A Fusion Architecture for Tracking a Group of People using a Distributed Sensor Network

Thyagaraju Damarla
U.S. Army Research Laboratory
Adelphi, MD 20783, U.S.A.
Email: thyagaraju.damarla.civ@mail.mil

Lance M. Kaplan
U.S. Army Research Laboratory
Adelphi, MD 20783, U.S.A.
Email: lance.m.kaplan.civ@mail.mil

***Abstract*—This paper presents an architecture to fuse the information from different spatially distributed unattended ground sensors (UGSs) consisting of multiple modalities to track a group of people walking along trails in open areas. The UGS system we used consists of ultrasonic, seismic and passive infrared (PIR) sensors. Estimating the size of the group is done by counting the number of targets using ultrasonic sensors. Determining the composition of the group is done using several classifiers. The fusion is done at the UGS level to fuse information from all the modalities to determine the presence of a target(s) and the information from multiple UGS systems are fused to track a group. The algorithms are tested on data collected in the wilderness with three UGS systems separated 75 m apart on a trail.**

Keywords: Personnel detection, decentralized fusion, tracking, non-negative matrix factorization, classification, basis vectors.

I. Introduction

There are number of reasons to find and track a group of people in the wilderness. One is to know their whereabouts and trace out the path they traveled in the event of their disappearance. Another reason is to protect long borders between two countries. Border patrol personnel are tasked with tracking down and intercepting people crossing the border illegally. They would rather intercept a larger group than a smaller one. In order to track a group of people, a number of unattended ground sensors (UGSs) are deployed along the known trails frequented by people crossing the border illegally. The current UGS systems available commercially [1] employ acoustic, seismic, PIR and magnetic sensors. The sensor algorithms detect people and notify the authorities of their detections using wireless communications links. Based on the sensor reports, the authorities use radar and imaging sensors to track the trespassers and make a decision as to intercept them or not depending on the size of the group and availability of personnel. A majority of the false alarms are caused by animals roaming in the wild. Based on the sensor reports from various sensors on different trails, the probability that a person would take a certain path to cross the border at a certain location is estimated [2] depending on the previous usage of various segments of the trails.

The UGS system considered in this paper for the purpose of tracking people is equipped with seismic, PIR and ultrasonic sensors. The ultrasonic sensors provide additional tracking information as well as the added capability to identify the targets at closer ranges. It is also assumed that the UGS is equipped with electronics to communicate to its immediate neighbor.

Whenever a group of targets (targets could be animals or people or combination of both) pass by the UGS, the sensors within the UGS analyze the data to determine the type of target and count the number of targets. Details of the analysis of each sensor data are presented in the section III. Individual UGS system may be able to determine the type and the number of targets in a group; however, in order to track the group, the information should be available to the other UGSs. The information should include the type of target, number of targets, etc. Moreover, in order to track a group, it is necessary to know the direction the group is traveling and also identify the group as being the same group that passed by the UGS down the path. Since, only non-imaging sensors are used, if two groups consist of the same number of targets, they are said to be the same group. Justification for such an assumption is that one does not have a preference for one group over the other as long as the size of the group is the same.

The organization of the paper is as follows. Section II presents the fusion architecture used to track a group of people. Section III presents the data analysis pertaining to classification and counting of the targets. Section III also presents the algorithms for fusion of distributed sensor data at the UGS level and fusion of information from different UGSs. In section IV, data collection in remote area is presented and also the results from the fusion algorithms. Conclusions are presented in section V.

II. Fusion Architecture used for Tracking People

As mentioned in the introduction, the majority of people walking in the wilderness use trails. These trails may or may not intersect. A web of trails is shown in Figure 1. The "dots" on the trails show possible deployments for the UGS. A majority of the times, people travel from the border to a destination or vice versa.

There are several questions that must be answered in order to track a group. They are: (a) what is the composition of the group, that is, how many different types of targets (people and animals) are in the group, (b) what is the direction of their

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
JUL 2013		**00-00-2013 to 00-00-2013**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
A Fusion Architecture for Tracking a Group of People Using a Distributed Sensor Network	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **U.S. Army Research Laboratory,Adelphi,MD,20783**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
Presented at the 16th International Conference on Information Fusion held in Istanbul, Turkey on 9-12 July 2013. Sponsored in part by Office of Naval Research Global.

14. ABSTRACT
This paper presents an architecture to fuse the information from different spatially distributed unattended ground sensors (UGSs) consisting of multiple modalities to track a group of people walking along trails in open areas. The UGS system we used consists of ultrasonic, seismic and passive infrared (PIR) sensors. Estimating the size of the group is done by counting the number of targets using ultrasonic sensors. Determining the composition of the group is done using several classifiers. The fusion is done at the UGS level to fuse information from all the modalities to determine the presence of a target(s) and the information from multiple UGS systems are fused to track a group. The algorithms are tested on data collected in the wilderness with three UGS systems separated 75 m apart on a trail.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**8**	

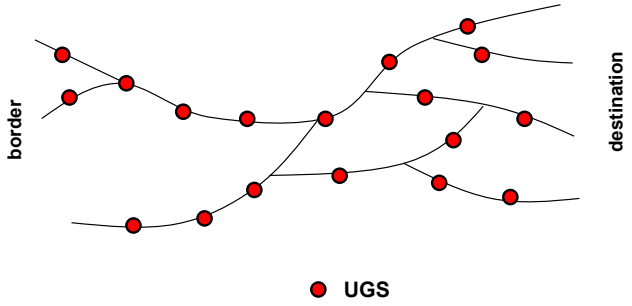

Figure 1. Labyrinth of trails

travel, and (c) is the group the current UGS system is detecting the same as the group who just passed by the previous UGS system. The answers to these questions have to be derived from the sensor data. As mentioned earlier, each UGS system consists of seismic, PIR, and ultrasonic sensors and their data are processed to obtain the required information. The analysis of the sensor data is presented in section III. We make some assumptions in order to answer the above questions:

- If somebody is riding an animal, the animal and the rider are considered as one target and categorized as an animal, since the rider does not contribute to the seismic signals and contributes very little to PIR and ultrasonic signatures. This will be clear in section III.
- Each target walks in a single file – this is a reasonable assumptions as the trails are narrow causing people walk one after the other. Moreover, a group of people bunched together is easier to detect by the border patrol than people who are spread out.
- Information from the neighboring UGS system is available before the current UGS data are processed.
- Each UGS system would determine the following information and pass it to its neighbors:
  - Direction of travel
  - Number of targets of each type

Since the sensors are distributed spatially over a large area, and the data from each UGS are available sequentially, the appropriate architecture for tracking people is the tandem architecture shown in Figure 2. When a trail branches into two, both branches are tested to determine the activity on the branch. The details of the fusion algorithm are presented in section III-D. In the next section, we present the analysis

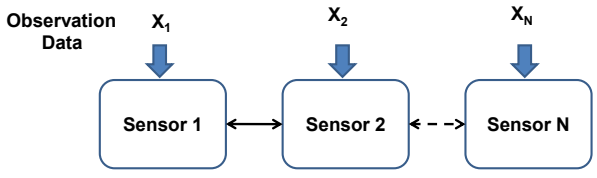

Figure 2. Architecture for tracking people

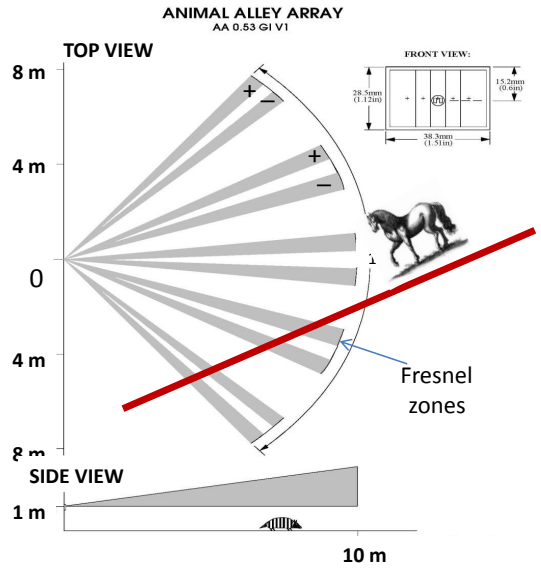

Figure 3. PIR sensor Fresnel zones

of each sensor's data to obtain the relevant information for tracking.

III. Sensor Data Analysis

Although, several sensor modalities, namely, acoustic, seismic, PIR, magnetic, electrostatic, micro-RADAR, etc., are used to detect people and animals [3], [4], in this paper we consider PIR, seismic, and ultrasonic sensors. Each one of these sensors brings a specific attribute to classify the targets and help in counting their number. Although we consider horses in this paper, the techniques developed here can be applied to other animals also. The data analysis is presented in the following sections.

A. Passive Infrared Sensor Data Analysis

PIR sensors are cheap and readily available and are used by most of the UGS systems [1] in the field today. PIR sensors are also used by homeowners and others as motion sensors to alert them of intruders. PIR sensors register the thermal radiation emitted by targets and objects in its field of view. The PIR transducer used here has dual element pyroelectric sensors connected back to back to cancel out the background radiation. As a result, the sensor output is proportionate to the thermal radiation of the target alone. The PIR sensor is also equipped with a Fresnel lens array in front of it dividing the field of view in to several zones as shown in Figure 3. Each zone is split in to positive and negative beams corresponding to the two elements of the PIR sensor. As the target intercepts the positive beam, a positive voltage is generated at the output of the sensor and similarly, a negative voltage is generated if the target passes through the negative beam. If the target occupies both the beams, the net output voltage would be zero.

Figure 4. Typical trail

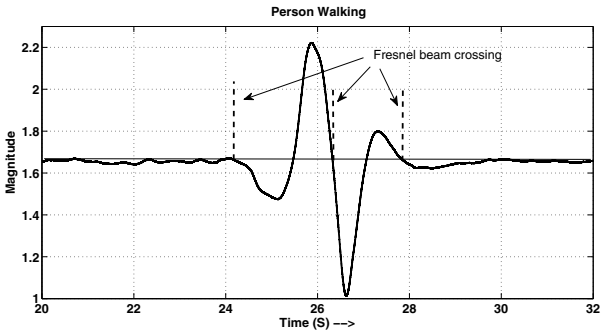

Figure 5. PIR signal produced by a person walking

A majority of the times the people crossing the border walk on the trails so as to cover a good distance in a short time. These trails are in general narrow and go through bushes. Placing the UGS systems along the trails gives an opportunity to get various sensor signatures at close range with good fidelity. One typical trail used for data collection is shown in Figure 4. The sensors are placed at the edges of the path with the Fresnel zones of the PIR sensor looking down the path. This placement allows a target to cut through several Fresnel zones. Depending on the length/width of the target, several zones are occupied by the target at a time. For example, a horse walking on the trail can touch two to three zones simultaneously. Figures 5 and 6 show the signals captured when a person and a horse walked on the path, respectively. From Figure 5, it is seen that as the person walks across the zones, the polarity of the signal changes from '$-$' to '$+$'

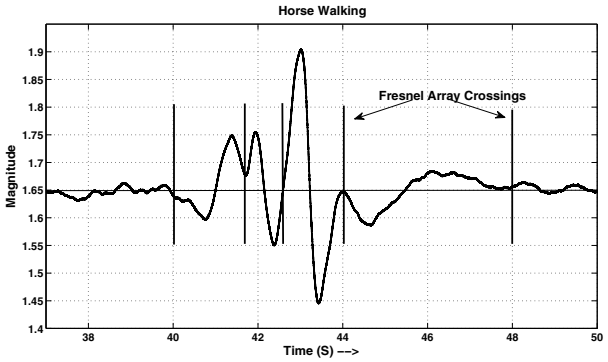

Figure 6. PIR signal produced by a horse walking

repeatedly. When a horse walks across the zones, the slope of the signal changes midstream and the polarity in Figure 6 is '$-++-+--+$' with respect to the average value of the signal. This is due to the fact that the horse is ≈ 6 ft in length and crosses several zones simultaneously. Moreover, the time it takes a person to cross all the beams is less compared to the time taken by a horse. Based on these observations an algorithm to discriminate people from animals is developed.

Algorithm 1: Target discrimination algorithm for a PIR sensor

- Estimate the duration 'τ_d' of the PIR signal activity.
- Determine if $\tau_d > T_h$, where T_h is the threshold. Note that T_h is different for each sensor placed on the trail which depends on the distance between the sensor and the trail and the angle at which the sensor's Fresnel beams intersect the trail. To have the same T_h for all sensors, the sensors should be placed at a predetermined distance from the trail and at a fixed angle with respect to the path.
- Determine if the polarity of the signal is '$--$' or '$++$' during the course of travel.
- If the polarity is '$--$' or '$++$' and $\tau > T_h$, declare the target is an animal, otherwise the target is a person.

B. Seismic sensor data analysis

Traditionally, seismic data analysis implies cadence analysis [5]. In this section, seismic data analysis is done to extract the basis vectors for each type of target so that they can be used to classify the targets. Extraction of basis vectors is done using non-negative matrix factorization (NMF) [6]–[9]. A brief description of NMF is presented here for the sake of continuity.

Let $[X]$ be a $[t \times \omega]$ matrix representing the short-time Fourier transform (STFT) with $X_{t,\omega}$ denoting an individual element of $[X]$ with variables in time t and frequency ω. NMF was first introduced by Lee and Seung [6], [7] and was adopted by others to minimize the cost function

$$\frac{1}{2}\sum_{t,\,\omega}\left|\,|X_{t,\,\omega}| - \sum_{i=1}^{k} W_{t,\,i}\,H_{i,\,\omega}\right|^2 + \lambda \sum_{t,\,i}|W_{t,\,i}|^1\,; \qquad (1)$$

$$s.t.\;\; H,\; W \geq 0,$$

where $[W]_{t\times k}$ and $[H]_{k\times\omega}$ are the weight and basis matrices, λ controls the sparsity on the weights and $|\bullet|^1$ denotes the L-1 norm. The number of basis vectors "k", which is also the rank of factorization, is usually chosen [6] so that $(\omega + t)k < \omega\ t$, and the product of W and H can be considered the compressed form of the original data matrix $[X]$. The NMF technique [9] takes a given set of observed vectors (rows of $[X]$) and uses them to find a set of basis vectors (rows of $[H]$) such that any observation can be represented as a linear combination of these basis vectors. NMF selects the recurrent patterns in the observations as the basis vectors. Note that the rows of $[H]$ are not orthogonal. Figure 7 shows a matrix factorization.

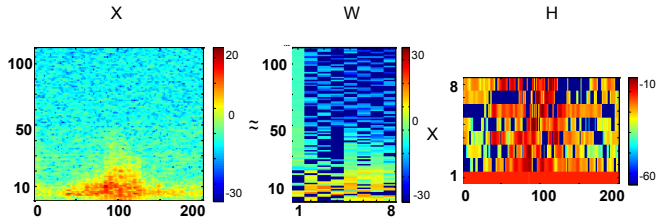

Figure 7. Factorization of matrix

The additive model in equation 1 implicitly assumes additivity of the magnitude spectra, which is only true if the phases are identical [9], [10] as shown below:

$$|X_{t,\omega}|\, e^{j\phi_0} = |C_{t,w}|\, e^{j\phi_1} + |D_{t,w}|\, e^{j\phi_2}$$

$$|X_{t,\omega}| = |C_{t,w}| + |D_{t,w}|\,; \;\; \text{iff} \;\; \phi_0 = \phi_1 = \phi_2.$$

To overcome this problem, Kameoka, et al. [9] developed a complex NMF technique by multiplying each spectral basis element by the $e^{j\phi}$ that best fits the signal:

$$\frac{1}{2}\sum_{t,\,\omega}\left|X_{t,\,\omega} - \sum_{i=1}^{k} W_{t,\,i}\; H_{i,\,\omega}\; e^{j\phi_{i,\omega,t}}\right|^2 + \lambda \sum_{t,\,i} |W_{t,\,i}|^1\,; \tag{2}$$

$$s.t. \;\; H,\; W \geq 0,$$

To overcome the problem of estimating the phase ϕ, we use discrete cosine transform (DCT) instead of STFT in this paper. We use both the positive and negative coefficients of DCT for basis vectors. When these basis vectors are used to represent a given signal, the phase is automatically estimated as is evident from the formulation of the problem given below. Let

$$X_i = \text{dct}\,(x_i(t))$$

be the DCT of the signal $x_i(t)$. In our case, the data are collected at 10 k samples per second, hence the DCT of $x_i(t)$ would have 10 k coefficients. However, the seismic sensor used has a frequency response of 0-250 Hz. As a result, only the first 500 out of 10 k coefficients carry most of the energy and are sufficient to reconstruct the time domain signal with negligible distortion.

Let X_i^+ and X_i^- denote the magnitudes of the positive and negative DCT coefficients such that $X_i = X_i^+ - X_i^-$. Let the matrix $[X_p] = \text{dct}(x_i); \forall i$ is the set of DCT coefficients for all the training data corresponding to people. Then the matrix $[X_p]$ can be written as

$$[X_p] = \left[X_p^+\right] - \left[X_p^-\right],$$

where $\left[X_p^+\right]$ representing the positive and $\left[X_p^-\right]$ representing the negative DCT coefficients of $[X_p]$, respectively. Similarly, $[X_a] = [X_a^+] - [X_a^-]$, represents the set for animals.

After matrix factorization using NMF, we get

$$\begin{aligned} \left[X_p^+\right] &\approx W_p\, H_p; \qquad \left[X_p^-\right] \approx \mathcal{W}_p\, \mathcal{H}_p \\ [X_a^+] &\approx W_a\, H_a; \qquad [X_a^-] \approx \mathcal{W}_a\, \mathcal{H}_a \end{aligned} \tag{3}$$

The matrices W and $\mathcal{W}$ represent the weight matrices, and the matrices H and $\mathcal{H}$ correspond to the basis vectors. Each row in H is a basis vector. As mentioned earlier, NMF selects recurrent patterns in the observations [9] as the rows of H. In other words, NMF projects all signals with same spectral content into single basis vector. Thus, we can represent a variety of seismic signals using a compact set of basis vectors.

Once the basis matrices H_p and $\mathcal{H}_p$ are available, the DCT coefficients X_t of a test signal $x(t)$ can be represented as a weighted sum of the basis vectors. The algorithm to estimate the weights and bases (subset of H and $\mathcal{H}$) is given below:

Algorithm 2: Algorithm to reconstruct target signatures

- Step 1: Normalize the test signal $x(t)$ after removing the mean. Compute $X_t = \text{dct}\,(x(t))$ and given by
$$X_t = X_t^+ - X_t^-$$
- Step 2: Estimate the weights $\omega = \{\omega_1, \omega_2, \cdots, \omega_k\}$ and $v = \{v_1, v_2, \cdots, v_k\}$ such that
$$\left|X_t^+ - \omega\; H\right|^2 + \left|X_t^- - v\; \mathcal{H}\right|^2; \quad \begin{array}{c} 0 \leq \omega_i, v_i \leq u_b; \\ \forall i \in \{1, 2, \cdots, k\} \end{array} \tag{4}$$
is minimum, u_b is typically 1. One may use any constrained nonlinear optimization program such as the "*fmincon*" function in MATLAB [11] to perform this task.
- Step 3: Non-zero weights ω and v give the bases used to represent X_t.
- Step 4: Reconstruct the signal $\hat{x}(t)$ by taking the inverse DCT of the difference $(\omega\; H \;-\; v\; \mathcal{H})$.

Notice we are estimating both X_t^+ and X_t^- in equation 4, that is, we are estimating the positive and negative DCT coefficients of a given signal hence the phase.

To determine whether the given X_t for an unknown signal belongs to human or animal we use combined basis vectors

$$H = \begin{bmatrix} H_p \\ H_a \end{bmatrix} \quad \mathcal{H} = \begin{bmatrix} \mathcal{H}_p \\ \mathcal{H}_a \end{bmatrix}$$

in the above algorithm 2. If the signal $x(t)$ belongs to people, the weights corresponding to the H_p and $\mathcal{H}_p$ will be higher compared to the weights belonging to H_a and $\mathcal{H}_a$.

Let S_p and S_a denote the total sum of the weights corresponding to the basis vectors for people and animals, respectively. If $x(t)$ is the seismic signature of a person walking alone, then $S_p > S_a$, since weights corresponding to animal basis vectors will be mostly zero or close to zero. Similarly, if $x(t)$ contains a signature of only an animal walking, then $S_a > S_p$. Else, if $x(t)$ contains a mixture, then $S_p \approx S_a > 0$, since the weights are non-zero for both. However, we cannot say which is greater as it depends on the basis vectors and how well they represent the given signal. In Figure 8 we plot the results of the algorithm for the data collected when a person was walking. The values of S_p are plotted as '*' and S_a as 'o'. Figure 8 clearly shows that the values of S_p are higher than S_a the majority of the time, indicating that the majority of the signatures are of a person walking. Results for the data collected with only a horse walking are shown in Figure 9.

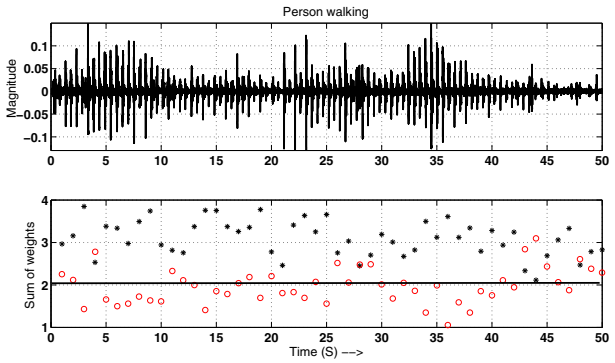

Figure 8. (a) Signature of a person walking (b) Classification

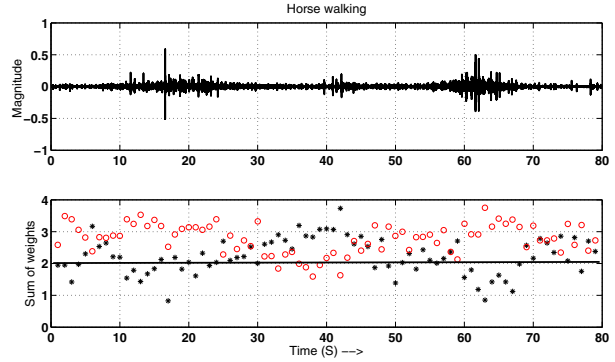

Figure 9. (a) Signature of a horse walking (b) Classification

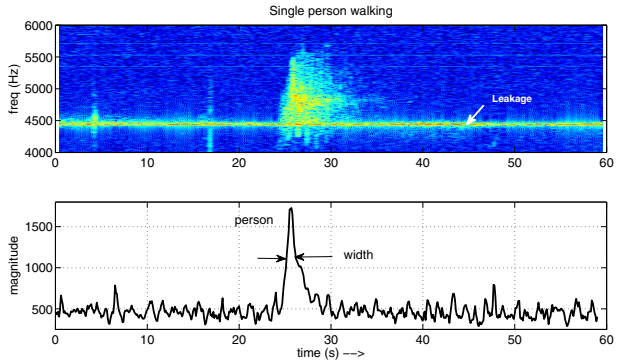

Figure 10. (a) Spectrogram of a person walking away from the sensor (b) Energy in the spectrum

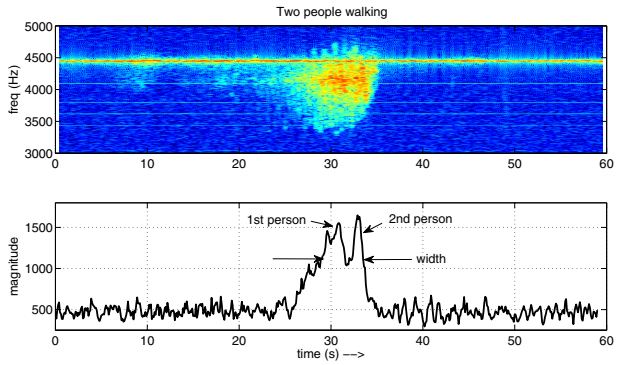

Figure 11. (a) Spectrogram of two people walking toward the sensor (b) Energy in the spectrum

C. *Ultrasonic Sensor Data Analysis*

In this section, we present ultrasonic data processing to identify the targets, estimate the direction of the targets traveling and count the number of targets. The ultrasonic transducer used has a transmitter that transmits a continuous wave at $f_c = 40$ kHz and a receiver to receive the Doppler shifted signals bounced off from the targets and other objects in the vicinity. The Doppler shift f_d is mainly due to torso and limb motion of the people and animals. For the sake of capturing the data, the return signals are down converted to 4.5 kHz by heterodyning the received signals with 44.5 kHz. Figures 10 and 11 show the spectrograms of one person and two people walking, respectively. In the first case, the person is walking away from the sensor. This can be seen in the spectrogram when the person comes into the beamwidth of the sensor, the signal strength is high initially and then as the person moves away from the sensor the signal strength decreases. In the second case, where two people are walking towards the sensor, the solid line around 4.5 kHz on both the spectrograms is due to leakage from the transmitter as well as returns from the stationary objects. It is well known that the Doppler frequency increases when the target is approaching the sensor. However, from Figure 11 the spectrum shows a decreasing frequency. In fact, the received frequency from the target is $f_c + f_d$, when it is subtracted from the 44.5 kHz during the heterodyning process, we get $4500 - f_d$, which is why increasing Doppler appears as decreasing frequency in the spectrogram. From Figures 10 and 11, we notice that the direction of travel can be inferred based on whether the spectral returns of the target are above 4500 Hz or below it.

Target Count: In order to count the number of targets, the energy in the spectrum is estimated and the data are plotted in Figures 10(b) and 11(b). The energy plot in Figure 11(b) shows two peaks indicating there are two targets. One can estimate the number of targets by counting the number of peaks with a minimum width 'ω' that is determined a priori.

Damarla, et. al. [12] presented different classification algorithms for animals and humans using ultrasonic signatures. When the signal-to-noise ratio (SNR) was high (> 6 db), they used the micro-Doppler characteristics exhibited by different targets to classify them. When the SNR levels were below 6 db, they used cepstral coefficients to classify the targets using a support vector machine (SVM) and multi-variate Gaussian (MVG) classifier. Details of the algorithms and the feature extraction can be found in [12]. Figure 12 shows the classification of ultrasonic signatures when a person and a horse are walked.

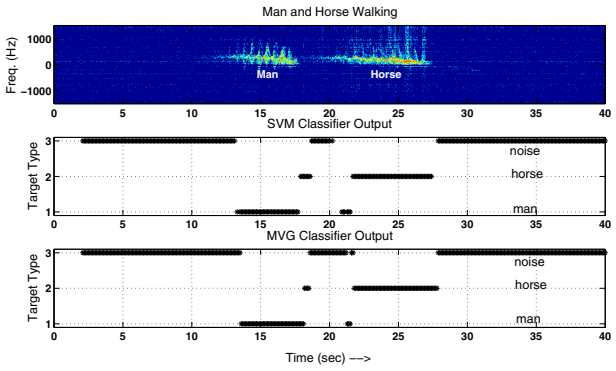

Figure 12. (a) Spectrogram of a person and horse walking, (b) classification results using SVM and (c) classification results using MVG.

D. Fusion of heterogeneous sensor data

In this section, we present the fusion algorithm used to fuse the data from the PIR, seismic and ultrasonic detectors. The same algorithm can be used to fuse the output of multiple UGSs also. The classification of the multimodal sensor data is done at each UGS system and the fusion is done at each UGS. Fusion of the multiple UGSs information is presented in section IV-A. It is assumed that each sensor receives an input

$$y_i = s_i + n_i,$$

where s_i is the signal and n_i is the noise. Each sensor makes a decision and generates a binary output

$$u_i = d(y_i) = \begin{cases} 1, & \text{if } \mathrm{H}_1 \text{ is true} \\ -1, & \text{if } \mathrm{H}_0 \text{ is true} \end{cases} \tag{5}$$

where H_1 and H_0 are the hypotheses for a target being present or absent, respectively, and "$d()$" is a decision function. The

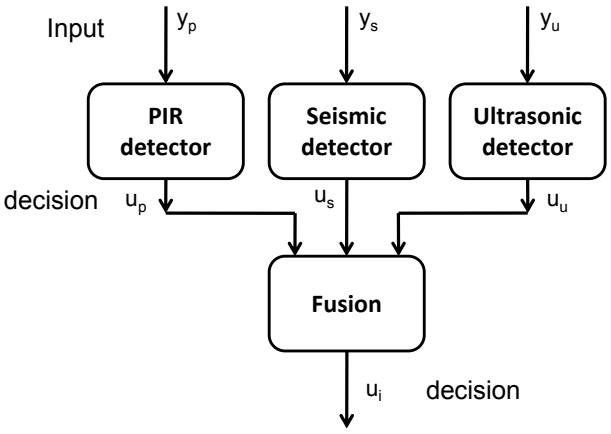

Figure 13. Fusion structure

fusion architecture used to fuse the sensor data at each UGS is shown in Figure 13. Each sensor makes a decision and passes its output to the fusion box, which makes the final decision as to the presence of a target or not. Chair and Varshney [13] presented an optimal decision rule for the likelihood ratio test (LRT)

$$\frac{P(u_1, \cdots, u_n|H_1)}{P(u_1, \cdots, u_n|H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \frac{P_0\ (C_{10} - C_{00})}{P_1\ (C_{01} - C_{11})} \tag{6}$$

where P_0 and P_1 are the priors for the two hypotheses and C_{ij} is the cost for selecting H_i when H_j is true for $i, j \in \{0, 1\}$. However, this approach requires the knowledge of the a-priori probabilities P_0 and P_1 which are hard to determine. An optimal fusion decision scheme using a Neyman-Pearson (N-P) test [14] mitigates the requirement on priors. The LRT can be rewritten as

$$\Lambda(u) = \frac{P(u_1, \cdots, u_N|H_1)}{P(u_1, \cdots, u_N|H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} t, \tag{7}$$

where "t" is the threshold determined by the desirable probability of false alarm. With the assumption that the decisions from the sensors are independent, we get

$$\Lambda(u) = \prod_{i=1}^{N} \frac{P(u_i|H_1)}{P(u_i|H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} t \tag{8}$$

The probability of false alarm P_F^f and probability of detection P_D^f at the fusion center are given by

$$\sum_{\Lambda(u) > t^*} P(\Lambda(u)|H_0) = P_F^f, \tag{9}$$

$$\sum_{\Lambda(u) > t^*} P(\Lambda(u)|H_1) = P_D^f, \tag{10}$$

where t^* is the threshold chosen to satisfy equation 9 for a desired P_F^f. The goal is to achieve with the N-P test, a pair (P_F^f, P_D^f) such that

$$P_F^f \leq \min_{i \in N} \{P_{F_i}\} \quad \text{and} \quad P_D^f > \min_{i \in N} \{P_{D_i}\} \tag{11}$$

where (P_{D_i}, P_{F_i}) is the N-P test level for sensor i, $i = 1, \cdots, N$. It is shown in [14] that, if $P_{D_i} = P_{D_j} = P_D$ and $P_{F_i} = P_{F_j} = P_F$ for all $i \neq j$, $i, j \in \{1, \cdots, N\}$, then the final probability of false alarm and detection are given by:

$$P_F^f = \sum_{k=|t_f^*|}^{N} \binom{N}{k} P_F^k \left(1 - P_F\right)^{N-k} \tag{12}$$

$$P_D^f = \sum_{k=|t_f^*|}^{N} \binom{N}{k} P_D^k \left(1 - P_D\right)^{N-k} \tag{13}$$

where $|t_f^*|$ indicates the smallest integer (number of sensors) required to exceed the threshold t^* to achieve the desired P_F^f. If the sensors have different probabilities of detection and false alarms, that is, $P_{D_i} \neq P_{D_j}$ and $P_{F_i} \neq P_{F_j}$, one can use the techniques described in [14] to estimate the final probabilities and the algorithm is given below:

Algorithm 2: Estimation of final probabilities for

sensors with different receiver operating curves (ROCs): Let the set $S = \{1, 2, \cdots, N\}$ represent all the sensors where N is the number of sensors, and $|t_f^*|$ is determined based on the final P_F desired. Then, let us denote all the subsets of S with k elements by S_k, where $|S_k| = \begin{pmatrix} N \\ k \end{pmatrix}$, then the final probability of false alarms and the final probability of detection are given by

$$P_F^f = \sum_{k=|t_f^*|}^{N} \left(\sum_{j=1}^{|S_k|} \left(\prod_{q=1}^{|s_j|;s_j \in S_k} P_{F_q} \prod_{r=1}^{|\bar{s}_j|} (1 - P_{F_r}) \right) \right) \quad (14)$$

$$P_D^f = \sum_{k=|t_f^*|}^{N} \left(\sum_{j=1}^{|S_k|} \left(\prod_{q=1}^{|s_j|;s_j \in S_k} P_{D_q} \prod_{r=1}^{|\bar{s}_j|} (1 - P_{D_r}) \right) \right) \quad (15)$$

where $\bar{s}$ is the complement of s.

IV. Data Collection and Algorithm Results

In order to verify the classification and fusion algorithms, we went to the southwest border of the USA to collect data. The data are collected on several terrains over a period of 5 days. On each day, several scenarios were enacted to collect the data. Some of the scenarios are: (a) one person, two people and three people walking, (b) one animal, two animals and three animals walking, (c) one animal and one person walking, and several people and several animals walking, etc. A total of 36 scenarios were enacted each day. One of the trails used for the data collection is shown in Figure 4. As mentioned earlier, we collected the data using three UGS systems, each system has (a) PIR sensor, (b) seismic sensor and (c) ultrasonic sensor. The separation between two UGS systems is about 60 to 75 m.

In order to estimate the overall probability of detection at each UGS system, we use the following algorithm:

Algorithm 3: Algorithm for computing probability of detection at a UGS

- Compute the probability of detection P_{D_i} for each sensor modality, namely, PIR, seismic and ultrasonic and their probabilities of false alarm P_{F_i}.
- The P_{D_i} and P_{F_i} are computed as follows:
  - Generate a training set for a sensor
  - Train the PIR, seismic and ultrasonic detectors using their respective training sets
  - Use individual detectors to test the entire data collected on all 5 days
  - Estimate the probability of false alarm based on the number of non-targets identified as targets and the probability of detection based on the number of targets correctly identified as targets.
- Perform the detection of new data using the detectors to obtain u_i, for all i.
- Estimate the probability of detection $P_F^{U_j}$ and $P_D^{U_j}$ for all UGS U_j, using equations 14 and 15 respectively.

Table I shows the P_{D_i} and P_{F_i} for the three sensors, and the overall probability of detection and false alarms for various $\left|t_f^*\right|$ values are given in Table II. From Table II, we notice that an improved overall probability of false alarm is achieved for $\left|t_f^*\right| \geq 2$, which implies that two or more sensors should have correct detection and classification.

PIR sensor	Seismic Sensor	Ultrasonic sensor
$P_F = 0.1$	$P_F = 0.1$	$P_F = 0.07$
$P_D = 0.9$	$P_D = 0.9$	$P_D = 0.8$

Table I
PROB. OF FALSE ALARM AND DETECTION FOR VARIOUS SENSORS

$\left\|t_f^*\right\|$	P_F^f	P_D^f
1.0000	0.2467	0.9980
2.0000	0.0226	0.9540
3.0000	0.0007	0.6480

Table II
THRESHOLD VERSUS FINAL PROB. OF FALSE ALARM AND DETECTION AT EACH UGS U_j.

Algorithm 4: Algorithm for computing probability of correct classification at a UGS

- First detect the presence of a target (animal/person) using each sensor modality
- If a target is detected, determine whether it is an animal or a person. Determine the probability of person and probability of animal at each modality
- Use equations 14 and 15 to determine the overall P_F and P_D at each UGS system. Note that, probability of false alarm here is the probability of animal classified as a person.
- One may use 'OR' or 'AND' rules [14] to fuse the information from different modalities instead of the equations 14 and 15.

A. Tracking of people

Figure 14 shows the flow chart used to track people using distributed UGSs. Each UGS system determines the direction of the targets, number of targets and the $\left(P_D^{U_j}, P_F^{U_j}\right)$. The information is passed on to its neighbors along with the information on number of targets and direction of travel. Hence, only a few bytes of data is sent to the near by UGS as opposed to sending the raw data from each sensor to a central node for data fusion. For tracking, we use a sliding window that covers the adjacent three UGSs to detect and report. This approach makes sure that the tracking is robust and done with high confidence levels (low P_F^f). As shown in Figure 14, if all three UGSs determine that the target's direction is same and the number of targets are same, the information from all three UGS will be fused using equations 12 and 13. An example of the fused output of three UGS is given in Table III, which uses the probability of false alarm $P_F^{U_j} = 0.0226$ and probability

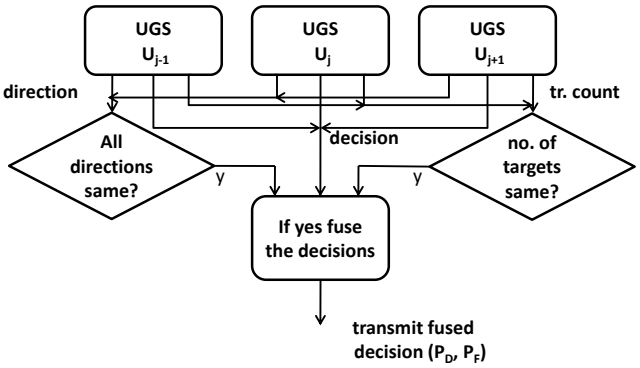

Figure 14. Flow chart used for tracking people using distributed UGS

of detection $P_D^{U_j} = 0.9540$ for all U_j (note that these values correspond to the 2^{nd} row in Table II).

$\lvert t_f^* \rvert$	P_F^f	P_D^f
1.0000	0.1426	0.9990
2.0000	0.0073	0.9720
3.0000	0.0001	0.7290

Table III
THRESHOLD VERSUS FINAL PROB. OF FALSE ALARM AND DETECTION AFTER FUSING INFORMATION FROM THREE UGS

V. CONCLUSION

In this paper, we considered the case of tracking a group of people (targets) walking along trails. We presented the algorithms to process individual sensor data for estimating the number of targets and their classification. We also showed that fusion of information of multimodal sensor at each UGS improves the overall probability of false alarm. Similarly, when we consider multiple UGSs to fuse their information, we improve the probability of detection and reduce the false alarm rate as well as our ability to track the group with high confidence.

Future efforts would include fusion of soft data that is obtained by scouts patrolling the border to predict the paths of travel of people crossing the border. Effort will be made to include the correlation between the observations of different modalities as the sensors are observing the same phenomenon and hence the assumption of independence is not quite valid.

REFERENCES

[1] *McQ, http://www.mcqinc.com/pdf/iScout_Datasheet-12-Jul2011.pdf.*

[2] G. Singh, K. G. Mehrotra, C. K. Mohan, and T. Damarla, "Inferring border crossing intentions with hidden markov models," in *Proceedings of the 24th international conference on Industrial engineering and other applications of applied intelligent systems conference on Modern approaches in applied intelligence - Volume Part I*, ser. IEA/AIE'11. Berlin, Heidelberg: Springer-Verlag, 2011, pp. 69–78. [Online]. Available: http://dl.acm.org/citation.cfm?id=2025756.2025767

[3] R. Damarla and D. Ufford, "Personnel detection using ground sensors," in *Proc. SPIE 6562*, vol. 6562, 2007, pp. 656 205–656 205–10. [Online]. Available: +http://dx.doi.org/10.1117/12.723212

[4] T. Damarla, A. Mehmood, and J. Sabatier, "Detection of people and animals using non-imaging sensors," in *Proceedings of the 14th International Conference on Information Fusion (FUSION)*, July 2011, pp. 1–8.

[5] K. M. Houston and D. P. McGaffigan, "Spectrum analysis techniques for personnel detection using seismic sensors," in *Proc. SPIE*, vol. 5090, Orlando, FL, 2003, pp. 162–173.

[6] D. D. Lee and H. Seung, "Learning the parts of objects by non-negative matrix factorization," in *Nature*, vol. 401, Oct 1999, pp. 788–791.

[7] ——, "Algorithm for non-negative matrix factorization," in *NIPS*, vol. 13, Vancouver, BC, Canada, 2001, pp. 556–562.

[8] A. Mehmood, T. Damarla, and J. Sabatier, "Separation of human and animal seismic signatures using non-negative matrix factorization," *Pattern Recognition Letters*, vol. 33, no. 16, pp. 2085–2093, June 2012. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0167865512002127

[9] H. Kameoka, N. Ono, K. Kashino, and S. Sagayama, "Complex nmf: A new sparse representation for acoustic signals," in *Acoustics Speech and Signal Processing (ICASSP), 2009 IEEE International Conference on*, 2009, pp. 3437–3440.

[10] B. King and L. Atlas, "Single-channel source separation using simplified-training complex matrix factorization," in *Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on*, Dallas, TX, march 2010, pp. 4206–4209.

[11] *MATLAB, http://www.mathworks.com*, July. date last viewed 07/15/2011.

[12] T. Damarla, M. Bradley, A. Mehmood, and J. Sabatier, "Classification of animals and people ultrasonic signatures," *Sensors Journal, IEEE*, vol. PP, no. 99, p. 1, 2012.

[13] Z. Chair and P. Varshney, "Optimal data fusion in multiple sensor detection systems," *IEEE Transaction on Aerospace and Electronic Systems*, vol. AES-22, no. 1, pp. 98–101, January 1986.

[14] S. Thomopoulos, R. Viswanathan, and D. Bougoulias, "Optimal decision fusion in multiple sensor systems," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. AES-23, no. 5, pp. 644 –653, sept. 1987.